



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2010

Translated Poisson approximation to equilibrium distributions of Markov population processes

Socoll, S N ; Barbour, A D

Abstract: The paper is concerned with the equilibrium distributions of continuous-time density dependent Markov processes on the integers. These distributions are known typically to be approximately normal, with $O(1/\sqrt{n})$ error as measured in Kolmogorov distance. Here, an approximation in the much stronger total variation norm is established, without any loss in the asymptotic order of accuracy; the approximating distribution is a translated Poisson distribution having the same variance and (almost) the same mean. Our arguments are based on the Stein–Chen method and Dynkin’s formula.

DOI: <https://doi.org/10.1007/s11009-009-9124-8>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-27116>

Journal Article

Published Version

Originally published at:

Socoll, S N; Barbour, A D (2010). Translated Poisson approximation to equilibrium distributions of Markov population processes. *Methodology and Computing in Applied Probability*, 12(4):567-586.

DOI: <https://doi.org/10.1007/s11009-009-9124-8>

Translated Poisson Approximation to Equilibrium Distributions of Markov Population Processes

Sanda N. Socoll · A. D. Barbour

Received: 15 February 2008 / Accepted: 11 February 2009 /
Published online: 7 March 2009
© Springer Science + Business Media, LLC 2009

Abstract The paper is concerned with the equilibrium distributions of continuous-time density dependent Markov processes on the integers. These distributions are known typically to be approximately normal, with $O(1/\sqrt{n})$ error as measured in Kolmogorov distance. Here, an approximation in the much stronger total variation norm is established, without any loss in the asymptotic order of accuracy; the approximating distribution is a translated Poisson distribution having the same variance and (almost) the same mean. Our arguments are based on the Stein–Chen method and Dynkin’s formula.

Keywords Continuous-time Markov process · Equilibrium distribution · Total-variation distance · Infinitesimal generator · Stein–Chen method · Point process

AMS 2000 Subject Classifications 60J75 · 62E17

1 Introduction

Density dependent Markov population processes, in which the transition rates depend on the density of individuals in the population, have proved widely useful as models in the social and life sciences: see, for example, the monograph of Kurtz (1981), in which approximations in terms of diffusions are extensively discussed, in

This work was supported in part by the Swiss National Foundation, Project No. 200020-107935/1; some of it was undertaken while the authors were visiting the Institute for Mathematical Sciences, National University of Singapore.

S. N. Socoll (✉) · A. D. Barbour
Institut für Mathematik, Universität Zürich-Irchel,
Winterthurerstr. 190, 8057 Zürich, Switzerland
e-mail: sanda.socoll@math.unizh.ch

the limit as the typical population size n tends to infinity. Here, we are interested in the behavior at equilibrium. Our starting point is the paper of Barbour (1980), in which conditions are given for the existence of an equilibrium distribution concentrated close to the deterministic equilibrium, together with a bound of order $O(1/\sqrt{n})$ on the Kolmogorov distance between the equilibrium distribution and a suitable normal distribution. We now show that this normal approximation can be substantially strengthened. Using a delicate argument based on the Stein–Chen method, we are able to establish an approximation in total variation in terms of a translated Poisson distribution. What is more, our error bounds with respect to this much stronger metric, and under weaker assumptions than those previously considered, are still of ideal order $O(1/\sqrt{n})$.

The first step in the argument is to establish the existence of an equilibrium distribution under suitable conditions, and to show that it is appropriately concentrated around the ‘deterministic’ equilibrium, defined to be the stationary point of an associated system of differential equations which describe the average drift of the process in the limit as $n \rightarrow \infty$; this is accomplished in Section 2. The closeness of this distribution to our approximation is then established in Section 4, by showing that Dynkin’s formula, applied in equilibrium, yields an equation not far removed from the Stein equation for a centred Poisson distribution, enabling ideas related to Stein’s method to be brought into play. An important element in obtaining an approximation in total variation is to show *a priori* that the equilibrium distribution is sufficiently smooth, in the sense that translating it by a single unit changes the distribution only by order $O(1/\sqrt{n})$ in total variation: see, for example, Röllin (2005). The corresponding argument is to be found in Section 3. We illustrate the results by applying them to a birth, death and immigration process, with births occurring in groups.

1.1 Basic Approach

We start by defining our density dependent sequence of Markov processes. For each $n \in \mathbb{N}$, let $Z_n(t)$, $t \geq 0$, be an irreducible continuous time pure jump Markov process taking values in \mathbb{Z} , with transition rates given by

$$i \rightarrow i + j \quad \text{at rate} \quad n\lambda_j\left(\frac{i}{n}\right), \quad i \in \mathbb{Z}, \quad j \in \mathbb{Z} \setminus \{0\},$$

where the $\lambda_j(\cdot)$ are prescribed functions on \mathbb{R} ; we set

$$z_n(t) := n^{-1}Z_n(t), \quad t \geq 0.$$

We then define an ‘average growth rate’ of the process z_n at $z \in n^{-1}\mathbb{Z}$ by

$$F(z) := \sum_{j \in \mathbb{Z} \setminus \{0\}} j\lambda_j(z),$$

and a ‘quadratic variation’ function by $n^{-1}\sigma^2(z)$, where

$$\sigma^2(z) = \sum_{j \in \mathbb{Z} \setminus \{0\}} j^2\lambda_j(z),$$

assumed to be finite for all $z \in \mathbb{R}$.

The ‘law of large numbers’ approximation shows that, for large n , the time dependent development of the process z_n runs close to the solution of the differential equation system $\dot{z} = F(z)$, with the same initial condition, and that there is an approximately diffusive behaviour on a scale $n^{-1/2}$ about this path (Kurtz 1970, 1971). If F has a single zero at a point c , and is such that c is globally attracting for the differential equation system, then Z_n has an equilibrium distribution Π_n that is approximately normal, and puts mass on a scale $n^{1/2}$ around nc (Barbour 1980). The corresponding asymptotic variance is given by $n^{1/2}v_c$ with $v_c := \frac{\sigma^2(c)}{-2F'(c)}$, provided that $F'(c) < 0$, and the error of the approximation in Kolmogorov distance is of ideal order $O(n^{-1/2})$ if only finitely many of the functions λ_j are non-zero.

In this paper, we strengthen this result, by proving an accurate approximation to the equilibrium distribution using another distribution on the integers. Under assumptions similar to those needed for the previous normal approximation, we prove that the distance in total variation between the centred equilibrium distribution $\Pi_n - [nc]$ and the centred Poisson distribution

$$\widehat{\text{Po}}(nv_c) := \text{Po}(nv_c) * \delta_{-[nv_c]}$$

is of order $O(n^{-1/2})$: here and subsequently, δ_r denotes the point mass on r , and $*$ denotes convolution. If infinitely many of the λ_j are allowed to be non-zero, but satisfy the analogue of a $(2 + \alpha)$ ’th moment condition, for some $0 < \alpha \leq 1$, we prove that the error is of order $O(n^{-\alpha/2})$.

The proof of our approximation runs as follows. The infinitesimal generator \mathcal{A}_n of Z_n , acting on a function h , is given by

$$(\mathcal{A}_n h)(i) := \sum_{j \in \mathbb{Z} \setminus \{0\}} n \lambda_j \left(\frac{i}{n} \right) [h(i+j) - h(i)], \quad i \in \mathbb{Z}.$$

In equilibrium, under appropriate assumptions on h , Dynkin’s formula implies that

$$\mathbb{E}(\mathcal{A}_n h)(Z_n) = 0. \quad (1.1)$$

The following lemma, whose proof we omit, expresses $\mathcal{A}_n h$ in an alternative form.

Lemma 1.1 *Suppose that $\sum_{j \in \mathbb{Z} \setminus \{0\}} j^2 \lambda_j(z) < \infty$ for all $z \in \mathbb{R}$. Then, for any function $h: \mathbb{Z} \rightarrow \mathbb{R}$ with bounded differences, we have*

$$(\mathcal{A}_n h)(i) = \frac{n}{2} \sigma^2 \left(\frac{i}{n} \right) \nabla g_h(i) + n F \left(\frac{i}{n} \right) g_h(i) + E_n(g, i), \quad (1.2)$$

where $\nabla f(i) := f(i) - f(i-1)$ and $g_h(i) := \nabla h(i+1)$ and, for any $i \in \mathbb{Z}$,

$$E_n(g, i) := -\frac{n}{2} F \left(\frac{i}{n} \right) \nabla g_h(i) + \sum_{j \geq 2} a_j(g, i) n \lambda_j \left(\frac{i}{n} \right) - \sum_{j \geq 2} b_j(g, i) n \lambda_{-j} \left(\frac{i}{n} \right), \quad (1.3)$$

with

$$2a_j(g, i) := -j(j-1) \nabla g(i) + 2 \sum_{k=1}^{j-1} k \nabla g(i+j-k) \quad (1.4)$$

$$= 2 \sum_{k=2}^j \binom{k}{2} \nabla^2 g_h(i+j-k+1); \quad (1.5)$$

$$2b_j(g, i) := j(j-1) \nabla g(i) - 2 \sum_{k=1}^{j-1} k \nabla g(i-j+k)$$

$$= 2 \sum_{k=2}^j \binom{k}{2} \nabla^2 g_h(i-j+k).$$

Writing Eq. 1.1 using the result of Lemma 1.1 leads to the required approximation, as follows. In equilibrium, Z_n/n is close to c , as is shown in the next section, and so the main part of Eq. 1.2 is close to

$$-F'(c) \left\{ \frac{n\sigma^2(c)}{-2F'(c)} \nabla g_h(i) - (i - nc)g_h(i) \right\},$$

because $F(c) = 0$. Here, the term in braces is very close to the Stein operator for the centred Poisson distribution $\widehat{\text{Po}}(nv_c)$ with $v_c = \frac{\sigma^2(c)}{-2F'(c)}$, applied to the function g_h : see Röllin (2005). Indeed, for any $v > 0$ and $B \subset \mathbf{Z}_v$, where $\mathbf{Z}_v := \{l \in \mathbb{Z}, l \geq -\lfloor v \rfloor\}$, one can write

$$\mathbb{1}_B(l) - \widehat{\text{Po}}(v)\{B\} = v \nabla g(l+1) - lg(l) + \langle v \rangle g(l), \quad l \in \mathbf{Z}_v, \quad (1.6)$$

for a function $g = g_{v,B}$ satisfying

$$\sup_{l \geq -\lfloor v \rfloor} |g(l+1)| \leq \min \left\{ 1, \frac{1}{\sqrt{v}} \right\}; \quad \sup_{l \geq -\lfloor v \rfloor} |\nabla g(l+1)| \leq \frac{1}{v}; \quad g(l) = 0, \quad l \leq -\lfloor v \rfloor, \quad (1.7)$$

where $\langle x \rangle := x - \lfloor x \rfloor$ denotes the fractional part of x ; note also, from Eqs. 1.6 and 1.7, that

$$\sup_l |lg(l)| \leq 3. \quad (1.8)$$

Replacing l in Eq. 1.6 by an integer valued random variable W then shows that, for any $B \subset \mathbf{Z}_v$,

$$\begin{aligned} & |\mathbb{P}[W \in B] - \widehat{\text{Po}}(v)\{B\}| \\ & \leq \sup_{g \in \mathcal{G}_v} |\mathbb{E}\{v \nabla g(W+1) - Wg(W) + \langle v \rangle g(W)\}| + \mathbb{P}[W < -\lfloor v \rfloor], \end{aligned} \quad (1.9)$$

where \mathcal{G}_v denotes the set of functions $g: \mathbb{Z} \rightarrow \mathbb{R}$ satisfying Eqs. 1.7 and 1.8. Hence, replacing W by Z_n and v by nv_c in Eq. 1.9, and comparing the expectation with Eq. 1.1 expressed using Lemma 1.1, the required approximation in total variation

can be deduced; for this part of the argument, we need in particular to show that, in equilibrium,

$$|\mathbb{E}\{\nabla g(Z_n + 1) - \nabla g(Z_n)\}| = |\mathbb{E}\{\nabla^2 g(Z_n + 1)\}| = O(n^{-3/2}), \quad (1.10)$$

and also that $\mathbb{E}|E_n(g, Z_n)| = O(n^{-\alpha/2})$ for any $g \in \mathcal{G}_{nv_c}$. The bound Eq. 1.10 follows from Corollary 3.3 in Section 3, and the latter estimate, which also uses Eq. 1.10, is the substance of Section 4.

1.2 Assumptions

We make the following assumptions on the functions λ_j . The first ensures that the deterministic differential equations have a unique equilibrium, which is sufficiently strongly attracting.

A1: There exists a unique c satisfying $F(c) = 0$; furthermore, $F'(c) < 0$ and, for any $\eta > 0$, $\mu_\eta := \inf_{|z-c| \geq \eta} |F(z)| > 0$.

The next assumption controls the global behaviour of the transition functions λ_j .

A2:

(a) For each $j \in \mathbb{Z} \setminus \{0\}$, there exists $c_j \geq 0$ such that

$$\lambda_j(z) \leq c_j(1 + |z - c|), \quad z \in \mathbb{R}, \quad (1.11)$$

where the c_j are such that, for some $0 < \alpha \leq 1$,

$$\sum_{j \in \mathbb{Z} \setminus \{0\}} |j|^{2+\alpha} c_j < \infty.$$

(b) For some $\lambda^0 > 0$,

$$\lambda_1(z) \geq 2\lambda^0, \quad z \in \mathbb{R}.$$

The moment condition on the c_j in Assumption A2(a) plays the same rôle as the analogous moment condition in the Lyapounov central limit theorem. Under this assumption, the ideal rate of convergence in the usual central limit approximation is the rate $O(n^{-\alpha/2})$ that we establish for our total variation approximation. Assumption A2(b) is important for establishing the smoothness of the equilibrium distribution Π_n . If, for instance, all jump sizes were multiples of 2, the approximation that we are concerned with would not be accurate in total variation.

We also require some assumptions concerning the local properties of the functions λ_j near c .

A3:

(a) There exist $\varepsilon > 0$ and $0 < \delta \leq 1$ and a set $J \subset \mathbb{Z} \setminus \{0\}$ such that

$$\inf_{|z-c| \leq \delta} \lambda_j(z) \geq \varepsilon \lambda_j(c) > 0, \quad j \in J;$$

$$\lambda_j(z) = 0 \quad \text{for all } |z - c| \leq \delta, \quad j \notin J.$$

(b) For each $j \in J$, λ_j is of class C^2 on $|z - c| \leq \delta$.

Assumptions A2 (a) and A3 imply in particular that the series $\sum_{j \in \mathbb{Z} \setminus \{0\}} j \lambda_j(z)$ and $\sum_{j \in \mathbb{Z} \setminus \{0\}} j^2 \lambda_j(z)$ are uniformly convergent on $|z - c| \leq \delta$, and that their sums, F and σ^2 respectively, are continuous there. They also imply that

$$\sum_{j \in \mathbb{Z} \setminus \{0\}} |j| n \lambda_j(i/n) = O(|i|), \quad |i| \rightarrow \infty,$$

so that the process Z_n is a.s. non-explosive, in view of Hamza and Klebaner (1995, Corollary 2.1).

The remaining assumptions control the derivatives of the functions λ_j near c .

A4: For δ as in A2,

$$L_1 := \sup_{j \in J} \frac{\|\lambda'_j\|_\delta}{\lambda_j(c)} < \infty,$$

where $\|f\|_\delta := \sup_{|z-c| \leq \delta} |f(z)|$.

This assumption implies in particular, in view of Assumptions A2–A3, that the series $\sum_{j \in \mathbb{Z} \setminus \{0\}} j \lambda'_j(z)$ and $\sum_{j \in \mathbb{Z} \setminus \{0\}} j^2 \lambda'_j(z)$ are uniformly convergent on $|z - c| \leq \delta$, that their sums are F' and $(\sigma^2)'$ respectively, and that F and σ^2 are of class C^1 on $|z - c| \leq \delta$.

A5: For δ as in A2,

$$L_2 := \sup_{j \in J} \frac{\|\lambda''_j\|_\delta}{|j| \lambda_j(c)} < \infty.$$

This assumption implies, in view of A2–A3, that the series $\sum_{j \in \mathbb{Z} \setminus \{0\}} j \lambda''_j(z)$ is uniformly convergent on $|z - c| \leq \delta$, its sum is F'' , and F is of class C^2 on $|z - c| \leq \delta$.

Our arguments make frequent use of the following theorem, which is a restatement in our setting of Hamza and Klebaner (1995, Theorem 3.2), and justifies Eq. 1.1.

Theorem 1.2 *Suppose that Z_n is non-explosive. Let h be a function satisfying*

$$(|\mathcal{A}_n| h)(i) := \sum_{j \in \mathbb{Z} \setminus \{0\}} \lambda_j \left(\frac{i}{n} \right) |h(i+j) - h(i)| \leq c_{n,h} (1 \vee |h(i)|), \quad |i| \rightarrow \infty, \quad (1.12)$$

for some $c_{n,h} < \infty$. Then, if $h(Z_n(0))$ is integrable, so is $h(Z_n(t))$ for any $t \geq 0$; moreover,

$$h(Z_n(t)) - h(Z_n(0)) - \int_0^t (\mathcal{A}_n h)(Z_n(s)) ds$$

is a martingale, and Dynkin's formula holds:

$$\mathbb{E}[h(Z_n(t)) - h(Z_n(0))] = \int_0^t \mathbb{E}(\mathcal{A}_n h)(Z_n(s)) ds. \quad (1.13)$$

2 Existence of the Equilibrium Distribution

In this section, we prove that Z_n has an equilibrium distribution which is suitably concentrated in the neighbourhood of nc .

Theorem 2.1 *Under Assumptions A1–A4, for all n large enough, Z_n has an equilibrium distribution Π_n , and*

$$\begin{aligned} \mathbb{E}_{\Pi_n} \{|z_n - c| \cdot \mathbf{1}(|z_n - c| > \delta)\} &= O(n^{-1}) \\ \mathbb{E}_{\Pi_n} \{(z_n - c)^2 \cdot \mathbf{1}(|z_n - c| \leq \delta)\} &= O(n^{-1}), \end{aligned} \quad (2.1)$$

for δ as in Assumption A3: here, as before, $z_n := n^{-1}Z_n$.

Proof The argument is based on suitable choices of Lyapounov functions. Consider the twice continuously differentiable function $V: \mathbb{R} \rightarrow \mathbb{R}_+$ defined by $V(z) := |z - c|^{2+\alpha}$, for the α in Assumption A2 (a). Since $V(c) = 0$ and $V(z) > 0$ for any $z \neq c$, and because

$$F(z)V'(z) = -|F(z)|(2+\alpha)|z - c|^{1+\alpha} < 0 \quad \text{for any } z \neq c, \quad (2.2)$$

while $F(c)V'(c) = 0$, we conclude that V is a Lyapounov function guaranteeing the asymptotic stability of the constant solution c of the equation $\dot{x} = F(x)$. We now use it to show the existence of Π_n .

Lemma 2.2 *Under the assumptions of Theorem 2.1, the function $h_V(i) := V(\frac{i}{n}) = |\frac{i}{n} - c|^{2+\alpha}$ fulfils the conditions of Theorem 1.2 with respect to the initial distribution δ_l , the point mass at l , for any $l \in \mathbb{Z}$.*

Proof Checking Eq. 1.12, we use Taylor approximation and Assumption A2 (a) to give

$$\begin{aligned} (|\mathcal{A}_n| h_V)(i) &\leq (2+\alpha)|z - c|^{1+\alpha} \sum_{j \in \mathbb{Z} \setminus \{0\}} |j| c_j (1 + |z - c|) \\ &\quad + \frac{(2+\alpha)(1+\alpha)|z - c|^\alpha}{2n} \sum_{j \in \mathbb{Z} \setminus \{0\}} j^2 c_j (1 + |z - c|) \end{aligned} \quad (2.3)$$

$$+ \frac{(2+\alpha)(1+\alpha)}{2n^{1+\alpha}} \sum_{j \in \mathbb{Z} \setminus \{0\}} |j|^{2+\alpha} c_j (1 + |z - c|), \quad (2.4)$$

where we write $z := i/n$. For $|z - c| < \delta \leq 1$, the estimate in Eq. 2.3 is uniformly bounded by

$$C_{1n} := 2(2+\alpha) \left\{ \sum_j |j| c_j + \frac{(1+\alpha)}{2n} \sum_j j^2 c_j + \frac{(1+\alpha)}{2n^{1+\alpha}} \sum_j |j|^{2+\alpha} c_j \right\} < \infty,$$

because of Assumption A2 (a); for $|z - c| \geq \delta$, we have the bound

$$(|\mathcal{A}_n| h_V)(i) \leq C_{1n} |z - c|^{2+\alpha} = C_{1n} h_V(i),$$

as required. \square

The above lemma allows us to apply Dynkin's formula to the function h_V . Using Taylor approximation as for Eq. 2.3, but now noting that the first order term

$$\sum_{j \in \mathbb{Z} \setminus \{0\}} \lambda_j(z) j V'(z) = F(z) V'(z)$$

can be evaluated using Eq. 2.2, it follows that

$$(\mathcal{A}_n h_V)(i) \leq -|F(z)|(2 + \alpha)|z - c|^{1+\alpha} + n^{-1} C_2 \leq n^{-1} C_2 \quad (2.5)$$

on $|z - c| \leq \delta$, for

$$C_2 = (2 + \alpha)(1 + \alpha) \left\{ \sum_j j^2 c_j + \sum_j |j|^{2+\alpha} c_j \right\} < \infty,$$

where, once again, $z := i/n$. On $|z - c| > \delta$ and under Assumption A2 (a), we have

$$\begin{aligned} (\mathcal{A}_n h_V)(i) &\leq -|F(z)|(2 + \alpha)|z - c|^{1+\alpha} \\ &\quad \times \left[1 - \frac{(1 + \alpha)}{2n|F(z)| \cdot |z - c|} \sum_{j \in \mathbb{Z} \setminus \{0\}} j^2 c_j (1 + |z - c|) \right. \\ &\quad \left. - \frac{(1 + \alpha)}{2n^{1+\alpha}|F(z)| \cdot |z - c|^{1+\alpha}} \sum_{j \in \mathbb{Z} \setminus \{0\}} |j|^{2+\alpha} c_j (1 + |z - c|) \right] \\ &\leq -\frac{\mu_\delta(2 + \alpha)}{2} |z - c|^{1+\alpha} \leq -\mu_\delta |z - c|^{1+\alpha}, \end{aligned} \quad (2.6)$$

as long as n is large enough that $n\delta \geq 1$ and

$$\frac{(1 + \delta)(1 + \alpha)}{n\delta\mu_\delta} \sum_{j \in \mathbb{Z} \setminus \{0\}} |j|^{2+\alpha} c_j < \frac{1}{2}.$$

Dynkin's formula 1.13 then implies, for such n , that

$$\begin{aligned} 0 \leq \mathbb{E}_i h_V(Z_n(t)) &= V(z) + \int_0^t \mathbb{E}_i(\mathcal{A}_n h_V)(Z_n(s)) ds \\ &\leq V(z) + \int_0^t \frac{C_2}{n} \mathbb{P}_i(|n^{-1} Z_n(s) - c| < \delta) ds \\ &\quad - \mu_\delta \int_0^t \mathbb{E}_i \{ |n^{-1} Z_n(s) - c|^{1+\alpha} \cdot \mathbb{1}(|n^{-1} Z_n(s) - c| \geq \delta) \} ds, \end{aligned}$$

for any $t > 0$ and $i \in \mathbb{Z}$, where \mathbb{P}_i and \mathbb{E}_i denote probability and expectation conditional on $Z_n(0) = i$. It now follows, for any $y \geq \delta$, that

$$\begin{aligned} & \frac{\mu_\delta y^{1+\alpha}}{t} \int_0^t \mathbb{P}_i(|n^{-1} Z_n(s) - c| \geq y) ds \\ & \leq \frac{\mu_\delta}{t} \int_0^t \mathbb{E}_i\{|n^{-1} Z_n(s) - c|^{1+\alpha} \cdot \mathbb{1}(|n^{-1} Z_n(s) - c| \geq y)\} ds \\ & \leq \frac{1}{t} V(z) + \frac{C_2}{nt} \int_0^t \mathbb{P}_i(|n^{-1} Z_n(s) - c| < \delta) ds, \end{aligned} \quad (2.7)$$

and, by letting $t \rightarrow \infty$, it follows that

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbb{P}_i(|n^{-1} Z_n(s) - c| \geq y) ds \leq \frac{C_2}{n\mu_\delta y^{1+\alpha}}.$$

This implies that a limiting equilibrium distribution Π_n for Z_n exists, see for instance Ethier and Kurtz (1986, Theorem 9.3, Chapter 4), and that, writing $z_n := n^{-1} Z_n$, we have

$$\mathbb{P}_{\Pi_n}(|z_n - c| \geq y) \leq \frac{C_2}{n\mu_\delta y^{1+\alpha}},$$

for any $y \geq \delta$. Furthermore,

$$\begin{aligned} \mathbb{E}_{\Pi_n}\{|z_n - c| \cdot \mathbb{1}(|z_n - c| \geq \delta)\} &= \int_\delta^\infty \mathbb{P}_{\Pi_n}(|z_n - c| \geq y) dy \\ &\leq \int_\delta^\infty \frac{C_2}{n\mu_\delta y^{1+\alpha}} dy = O(n^{-1}), \end{aligned}$$

proving the first inequality in Eq. 2.1.

For the second inequality in Eq. 2.1, we define a function $\tilde{V}: \mathbb{R} \rightarrow \mathbb{R}$, which is of class $C^2(\mathbb{R})$, is bounded and has uniformly bounded first and second derivatives on \mathbb{R} , fulfils the conditions of Theorem 1.2, and satisfies $F(z)\tilde{V}'(z) = -(z - c)^2$ on $|z - c| \leq \delta$.

In view of the latter property, we begin by letting $v: [c - \delta, c + \delta] \rightarrow \mathbb{R}_+$ be the function defined by

$$v(z) := \int_c^z \frac{-(x - c)^2}{F(x)} dx,$$

with $v(c) = 0$. Note that v is well defined, since $F'(x) < 0$ on a small enough neighborhood of c , by Assumptions A1 and A4, and that $v(z) > 0$ for any $z \neq c$. Furthermore, in view of Assumptions A1 and A4,

$$v'(z) = -\frac{(z - c)^2}{F(z)} \text{ and } v''(z) = \frac{(z - c)^2 F'(z) - 2(z - c)F(z)}{F^2(z)}$$

exist and are continuous on $|z - c| \leq \delta$, since $|F(z)| > 0$ for $z \neq c$, $F(z) \sim F'(c)(z - c)$ for $z \rightarrow c$, and F' is continuous. In particular, we have

$$v'(c) = \lim_{z \rightarrow c} v'(z) = 0 \text{ and } v''(c) = \lim_{z \rightarrow c} v''(z) = -\frac{1}{F'(c)} > 0. \quad (2.8)$$

Now define the function \tilde{V} to be identical with v on $|z - c| \leq \delta$, and continued in $z \leq c - \delta$ and in $z \geq c + \delta$ in such a way that the function is still C_2 , and takes the same fixed value everywhere on $|z - c| \geq 2\delta$. Let

$$C_3 := \max \left\{ \sup_{z \in \mathbb{R}} \tilde{V}(z), \sup_{z \in \mathbb{R}} |\tilde{V}'(z)|, \sup_{z \in \mathbb{R}} |\tilde{V}''(z)| \right\}.$$

Lemma 2.3 *Under the assumptions of Theorem 2.1, the function $\tilde{h}_V(i) := \tilde{V}(\frac{i}{n})$ fulfils the conditions of Theorem 1.2 with respect to the initial distribution Π_n .*

Proof Since $\tilde{h}_V(i)$ is bounded, it follows that $\mathbb{E}_{\Pi_n} |\tilde{h}_V(Z_n)| < \infty$. $|\mathcal{A}_n| \tilde{h}_V$ is also bounded, since, for $|n^{-1}i - c| \leq 4\delta$, by Assumption A2 (a),

$$\left(|\mathcal{A}_n| \tilde{h}_V \right) (i) \leq C_3 \sum_{j \in \mathbb{Z} \setminus \{0\}} c_j (1 + 4\delta),$$

while, for $|n^{-1}i - c| > 4\delta$,

$$\begin{aligned} \left(|\mathcal{A}_n| \tilde{h}_V \right) (i) &\leq C_3 \sum_{j: |j+i-nc| \leq 2n\delta} c_j (1 + |n^{-1}i - c|) \\ &\leq C_3 \left\{ \sum_{j \in \mathbb{Z} \setminus \{0\}} j c_j \right\} \frac{1 + |n^{-1}i - c|}{|i - nc| - 2n\delta} \leq C_3 \left\{ \sum_{j \in \mathbb{Z} \setminus \{0\}} j c_j \right\} \frac{1 + 4\delta}{2n\delta}. \end{aligned}$$

□

We now apply Dynkin's formula to \tilde{h}_V , obtaining

$$0 = \mathbb{E}_{\Pi_n} \{ (\mathcal{A}_n \tilde{h}_V)(Z_n) \} \leq \mathbb{E}_{\Pi_n} \left\{ F(Z_n) \tilde{V}'(Z_n) + \sum_{j \in \mathbb{Z} \setminus \{0\}} \lambda_j(Z_n) \frac{j^2}{2n} C_3 \right\}.$$

Hence it follows that

$$\begin{aligned} &\mathbb{E}_{\Pi_n} \left\{ -F(Z_n) \tilde{V}'(Z_n) \cdot \mathbb{1}(|Z_n - c| \leq \delta) \right\} \\ &\leq \mathbb{E}_{\Pi_n} \left\{ F(Z_n) \tilde{V}'(Z_n) \cdot \mathbb{1}(|Z_n - c| > \delta) + \sum_{j \in \mathbb{Z} \setminus \{0\}} \lambda_j(Z_n) \frac{j^2}{2n} C_3 \right\}, \end{aligned}$$

whence we obtain

$$\begin{aligned} &\mathbb{E}_{\Pi_n} \left\{ (Z_n - c)^2 \cdot \mathbb{1}(|Z_n - c| \leq \delta) \right\} \\ &\leq \mathbb{E}_{\Pi_n} \left\{ |F(Z_n) \tilde{V}'(Z_n)| \cdot \mathbb{1}(|Z_n - c| > \delta) \right\} + C_3 \mathbb{E}_{\Pi_n} \left\{ \sum_{j \in \mathbb{Z} \setminus \{0\}} \lambda_j(Z_n) \frac{j^2}{2n} \right\} \\ &\leq C_3 \sum_{j \in \mathbb{Z} \setminus \{0\}} \left(2|j| + \frac{j^2}{n} \right) c_j \mathbb{E}_{\Pi_n} \{ |Z_n - c| \cdot \mathbb{1}(|Z_n - c| > \delta) \} + \frac{C_3}{2n} \sup_{|z-c| \leq \delta} \sigma^2(z). \end{aligned}$$

Using the first inequality in Eq. 2.1 and Assumptions A2 and A3, we conclude that

$$\mathbb{E}_{\Pi_n} \{ (z_n - c)^2 \cdot \mathbb{1}(|z_n - c| \leq \delta) \} = O(n^{-1}),$$

proving the second inequality in Eq. 2.1. \square

Corollary 2.4 *Under Assumptions A1–A4,*

$$\mathbb{E}_{\Pi_n} \{ |z_n - c| \} = O(n^{-1/2}).$$

Proof Using Hölder’s inequality, we obtain

$$\begin{aligned} \mathbb{E}\{|z_n - c|\} &= \mathbb{E}_{\Pi_n} \{ |z_n - c| \cdot \mathbb{1}(|z_n - c| > \delta) \} + \mathbb{E}_{\Pi_n} \{ |z_n - c| \cdot \mathbb{1}(|z_n - c| \leq \delta) \} \\ &\leq \mathbb{E} \{ |z_n - c| \cdot \mathbb{1}(|z_n - c| > \delta) \} + \sqrt{\mathbb{E}_{\Pi_n} \{ (z_n - c)^2 \cdot \mathbb{1}(|z_n - c| \leq \delta) \}}. \end{aligned}$$

The corollary now follows from Theorem 2.1. \square

Corollary 2.5 *Under Assumptions A1–A4, for any $0 < \delta' \leq \delta$,*

$$\mathbb{P}_{\Pi_n} [|z_n - c| > \delta'] = O(n^{-1}).$$

Proof It follows from Chebyshev’s inequality and Theorem 2.1 that

$$\begin{aligned} \mathbb{P}_{\Pi_n} [|z_n - c| I[|z_n - c| \leq \delta] > \delta'/2] \\ \leq 4 \mathbb{E}_{\Pi_n} \{ |z_n - c|^2 I[|z_n - c| \leq \delta] \} / (\delta')^2 = O(n^{-1}), \end{aligned}$$

and that

$$\mathbb{P}_{\Pi_n} [|z_n - c| > \delta] \leq \mathbb{E}_{\Pi_n} \{ |z_n - c| I[|z_n - c| > \delta] \} / \delta = O(n^{-1}),$$

from which the corollary follows. \square

3 The Distance Between Π_n and Its Unit Translation

A key step in the argument leading to our approximation is to establish that the equilibrium distribution Π_n of Z_n is sufficiently smooth. In order to do so, we first need to prove an auxiliary result, showing that, if the process Z_n starts near enough to nc , then it remains close to nc with high probability over any finite time interval. This is the substance of the following lemma.

Lemma 3.1 *Under Assumptions A1–A4, for any $0 < \eta \leq \delta$, there exists a constant $K_{U,\eta} < \infty$ such that*

$$\mathbb{P} \left[\sup_{t \in [0, U]} |Z_n(t) - nc| > n\eta \mid Z_n(0) = i \right] \leq n^{-1} K_{U,\eta},$$

uniformly in $|i - nc| \leq n\eta e^{-K_1 U} / 2$, where $K_1 := \|F'\|_\delta$.

Proof It follows directly from Assumption A2 (a) that h defined by $h(j) = j$ satisfies condition Eq. 1.12. Fix $Z_n(0) = i$, and define

$$\tau_\eta := \inf\{t \geq 0: |Z_n(t) - nc| > n\eta\}. \quad (3.1)$$

Then it follows from Theorem 1.2 that

$$\mathcal{M}_n(t) := Z_n(t \wedge \tau_\eta) - i - \int_0^{t \wedge \tau_\eta} nF(z_n(s)) ds$$

is a martingale with expectation 0, and with expected quadratic variation no larger than

$$nt \sum_{j \in \mathbb{Z} \setminus \{0\}} j^2 c_j (1 + \eta) \quad (3.2)$$

at time t (see Hamza and Klebaner (1995, Corollary 3)); here, as earlier, $z_n := n^{-1} Z_n$. Hence we have

$$|z_n(t \wedge \tau_\eta) - c| \leq \frac{1}{n} \left\{ \sup_{s \in [0, U]} |\mathcal{M}_n(s)| + |i - nc| \right\} + \int_0^{t \wedge \tau_\eta} |F(z_n(s))| ds,$$

for any $0 \leq t \leq U$, and also, from Assumptions A1–A4, we have

$$|F(z)| = |F(z) - F(c)| \leq \sup_{|y-c| \leq \delta} |F'(y)| |z - c|.$$

Hence it follows that

$$\int_0^{t \wedge \tau_\eta} |F(z_n(s))| ds \leq K_1 \int_0^{t \wedge \tau_\eta} |z_n(s) - c| ds.$$

Gronwall's inequality now implies that

$$|z_n(t \wedge \tau_\eta) - c| \leq n^{-1} \left\{ \sup_{s \in [0, U]} |\mathcal{M}_n(s)| + |i - nc| \right\} e^{K_1 t},$$

for any $0 \leq t \leq U$, and so, for $|i - nc| \leq n\eta e^{-K_1 U}/2$,

$$\sup_{t \in [0, U]} |z_n(t \wedge \tau_\eta) - c| \leq \eta/2 + n^{-1} \sup_{s \in [0, U]} |\mathcal{M}_n(s)| e^{K_1 U}. \quad (3.3)$$

We have thus shown that

$$\mathbb{P} \left[\sup_{t \in [0, U]} |z_n(t) - c| > \eta \mid Z_n(0) = i \right] \leq \mathbb{P} \left[\sup_{s \in [0, U]} |\mathcal{M}_n(s)| > ne^{-K_1 U} \eta/2 \mid Z_n(0) = i \right]. \quad (3.4)$$

But by Kolomogorov's inequality, from Eq. 3.2, we have

$$\mathbb{P} \left[\sup_{s \in [0, U]} |\mathcal{M}_n(s)| > ne^{-K_1 U} \eta/2 \mid Z_n(0) = i \right] \leq 4n^{-1} \eta^{-2} e^{2K_1 U} U \sum_{j \in \mathbb{Z} \setminus \{0\}} j^2 c_j (1 + \eta), \quad (3.5)$$

completing the proof. \square

We can now prove the main theorem of this section.

Theorem 3.2 *Under Assumptions A1–A4, there exists a constant $K > 0$ such that*

$$d_{TV} \{ \Pi_n, \Pi_n * \delta_1 \} \leq K n^{-1/2},$$

where $\Pi_n * \delta_1$ denotes the equilibrium distribution Π_n of Z_n , translated by 1.

Proof Because we have little *a priori* information about Π_n , we fix any $U > 0$, and use the stationarity of Π_n to give the inequality

$$d_{TV} \{ \Pi_n, \Pi_n * \delta_1 \} \leq \sum_{i \in \mathbb{Z}} \Pi_n(i) d_{TV} \{ \mathcal{L}(Z_n(U) \mid Z_n(0) = i), \mathcal{L}(Z_n(U) + 1 \mid Z_n(0) = i) \}. \quad (3.6)$$

By Corollary 2.5, we thus have, for any $\delta' \leq \delta$,

$$d_{TV} \{ \Pi_n, \Pi_n * \delta_1 \} \leq D_{1n}(\delta') + O(n^{-1}), \quad (3.7)$$

where

$$D_{1n}(\delta') := \sum_{i: |i-nc| \leq \delta'} \Pi_n(i) d_{TV} \{ \mathcal{L}(Z_n(U) \mid Z_n(0) = i), \mathcal{L}(Z_n(U) + 1 \mid Z_n(0) = i) \}.$$

This alters our problem to one of finding a bound of similar form, but now involving the transition probabilities of the chain Z_n over a finite time U , and started in a fixed state i which is relatively close to nc .

We now use the fact that the upward jumps of length 1 occur at least as fast as a Poisson process of rate λ^0 , something that will be used to derive the smoothness that we require. We realize the chain Z_n with $Z_n(0) = i$ in the form $N_n + X_n$, for the bivariate chain (N_n, X_n) having transition rates

$$\begin{aligned} (l, m) &\rightarrow (l + 1, m) \text{ at rate } n\lambda^0 \\ (l, m) &\rightarrow (l, m + 1) \text{ at rate } n \left[\lambda_1 \left(\frac{l+m}{n} \right) - \lambda^0 \right] \\ (l, m) &\rightarrow (l, m + j) \text{ at rate } n\lambda_j \left(\frac{l+m}{n} \right), \text{ for any } j \in \mathbb{Z}, j \neq 0, 1, \end{aligned}$$

and starting at $(0, i)$. This allows us to deduce that

$$\begin{aligned} &d_{TV} \{ \mathcal{L}(Z_n(U) \mid Z_n(0) = i), \mathcal{L}(Z_n(U) + 1 \mid Z_n(0) = i) \} \\ &= \frac{1}{2} \sum_{k \in \mathbb{Z}} | \mathbb{P}(Z_n(U) = k \mid Z_n(0) = i) - \mathbb{P}(Z_n(U) = k - 1 \mid Z_n(0) = i) | \\ &= \frac{1}{2} \sum_{k \in \mathbb{Z}} \left| \sum_{l \geq 0} \mathbb{P}(N_n(U) = l) \mathbb{P}(X_n(U) = k - l \mid N_n(U) = l, X_n(0) = i) \right. \\ &\quad \left. - \sum_{l \geq 1} \mathbb{P}(N_n(U) = l - 1) \mathbb{P}(X_n(U) = k - l \mid N_n(U) = l - 1, X_n(0) = i) \right| \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{2} \sum_{k \in \mathbb{Z}} \sum_{l \geq 0} |\mathbb{P}(N_n(U) = l) - \mathbb{P}(N_n(U) = l-1)| f_{l,i}^U(k-l) \\ &\quad + \frac{1}{2} \sum_{k \in \mathbb{Z}} \sum_{l \geq 1} \mathbb{P}(N_n(U) = l-1) |f_{l,i}^U(k-l) - f_{l-1,i}^U(k-l)|, \end{aligned} \quad (3.8)$$

where

$$f_{l,i}^U(m) := \mathbb{P}(X_n(U) = m \mid N_n(U) = l, X_n(0) = i). \quad (3.9)$$

Since, from Barbour et al. (1992, Theorem 1.C),

$$\sum_{l \geq 0} |\mathbb{P}(N_n(U) = l) - \mathbb{P}(N_n(U) = l-1)| \leq \frac{1}{\sqrt{n\lambda^0 U}} = O\left(\frac{1}{\sqrt{n}}\right), \quad (3.10)$$

the first term in Eq. 3.8 is bounded by $1/\{\sqrt{n\lambda^0 U}\}$, yielding a contribution of the same size to $D_{1n}(\delta')$ in Eq. 3.7, and it remains only to control the differences between the conditional probabilities $f_{l,i}^U(m)$ and $f_{l-1,i}^U(m)$.

To make the comparison between $f_{l,i}^U(m)$ and $f_{l-1,i}^U(m)$, we first condition on the whole Poisson paths of N_n leading to the events $\{N_n(U) = l\}$ and $\{N_n(U) = l-1\}$, respectively, chosen to be suitably matched; we write

$$\begin{aligned} f_{l,i}^U(m) &= \frac{1}{U^l} \int_{[0,U]^l} ds_1 \dots ds_{l-1} ds^* \mathbb{P}(X_n(U) = m \mid \\ &\quad N_n[0, U] = v^l(\cdot; s_1, \dots, s_{l-1}, s^*), X_n(0) = i); \\ f_{l-1,i}^U(m) &= \frac{1}{U^l} \int_{[0,U]^l} ds_1 \dots ds_{l-1} ds^* \mathbb{P}(X_n(U) = m \mid \\ &\quad N_n[0, U] = v^{l-1}(\cdot; s_1, \dots, s_{l-1}), X_n(0) = i), \end{aligned} \quad (3.11)$$

where

$$v^r(u; t_1, \dots, t_r) := \sum_{i=1}^r \mathbb{1}_{[0,u]}(t_i),$$

and $Y[0, u]$ is used to denote $(Y(s), 0 \leq s \leq u)$. Fixing s_1, s_2, \dots, s_{l-1} , let \mathbb{P}_{i,s^*} denote the distribution of X_n conditional on $N_n[0, U] = v^l(\cdot; s_1, \dots, s_{l-1}, s^*)$ and $X_n(0) = i$, and let \mathbb{P}_i denote that conditional on $N_n[0, U] = v^{l-1}(\cdot; s_1, \dots, s_{l-1})$ and $X_n(0) = i$; let $\rho_{s^*}(u, x)$ denote the Radon–Nikodym derivative $d\mathbb{P}_{i,s^*}/d\mathbb{P}_i$ evaluated at the path $x[0, u]$. Then

$$\mathbb{P}_{i,s^*}[X_n(U) = m] = \int_{\{x[0,U]: x(U)=m\}} \rho_{s^*}(U, x) d\mathbb{P}_i(x[0, U]),$$

and hence

$$\mathbb{P}_{i,s^*}[X_n(U) = m] - \mathbb{P}_i[X_n(U) = m] = \int \mathbb{1}_{\{m\}}(x(U)) \{\rho_{s^*}(U, x) - 1\} d\mathbb{P}_i(x[0, U]). \quad (3.12)$$

Thus

$$\begin{aligned} & \sum_{m \in \mathbb{Z}} |f_{l,i}^U(m) - f_{l-1,i}^U(m)| \\ & \leq \frac{1}{U^l} \int_{[0,U]^l} ds_1 \dots ds_{l-1} ds^* \sum_{m \in \mathbb{Z}} \mathbb{E}_i \{ \mathbb{1}_{\{m\}}(X_n(U)) |\rho_{s^*}(U, X_n) - 1| \} \\ & \leq \frac{2}{U^l} \int_{[0,U]^l} ds_1 \dots ds_{l-1} ds^* \mathbb{E}_i \{ [1 - \rho_{s^*}(U, X_n)]_+ \}. \end{aligned} \quad (3.13)$$

To evaluate the expectation, note that $\rho_{s^*}(u, X_n)$, $u \geq 0$, is a \mathbb{P}_i -martingale with expectation 1. Now, if the path $x[0, U]$ has r jumps at times $t_1 < \dots < t_r$, writing

$$y(v) := x(v) + v^{l-1}(v; s_1, \dots, s_{l-1}), \quad y_k := y(t_k), \quad j_k := y_k - y_{k-1},$$

we have

$$\rho_{s^*}(u, x) = \begin{cases} 1 & \text{if } u < s^*; \\ \exp \left(-n \int_{s^*}^u \left\{ \hat{\lambda}(y(v) + n^{-1}) - \hat{\lambda}(y(v)) \right\} dv \right) \\ \times \prod_{\{k: s^* \leq t_k \leq u\}} \left\{ \hat{\lambda}_{j_k}(y_{k-1} + n^{-1}) / \hat{\lambda}_{j_k}(y_{k-1}) \right\} & \text{if } u \geq s^*, \end{cases}$$

where $\hat{\lambda}_j(\cdot) = \lambda_j(\cdot)$ if $j \neq 1$ and $\hat{\lambda}_1(\cdot) = \lambda_1(\cdot) - \lambda^0$, and where $\hat{\lambda}(\cdot) := \sum_{j \in \mathbb{Z} \setminus \{0\}} \hat{\lambda}_j(\cdot)$. Thus, in particular, $\rho_{s^*}(u, x)$ is absolutely continuous except for jumps at the times t_k . Then also, from Assumptions A3 (a) and A4,

$$\left| \frac{\lambda_j(y + n^{-1})}{\lambda_j(y)} - 1 \right| \leq \frac{\|\lambda'_j\|_\delta}{n\varepsilon \lambda_j(c)} \leq |j| L_1 / \{n\varepsilon\},$$

uniformly in $|y - c| \leq \delta$, for each $j \in J$. Hence it follows that, if we define the stopping times

$$\begin{aligned} \tau_\delta &:= \inf\{u \geq 0 : |X_n(u) + v^{l-1}(u; s_1, \dots, s_{l-1}) - nc| > n\delta\}; \\ \phi &:= \inf\{u \geq 0 : \rho_{s^*}(u, X_n) \geq 2\}, \end{aligned} \quad (3.14)$$

then the expected quadratic variation of the martingale $\rho_{s^*}(u, X_n)$ up to the time $\min\{U, \tau_\delta, \phi\}$ is at most

$$4U \sum_{j \in \mathbb{Z} \setminus \{0\}} \left(\frac{|j| L_1}{n\varepsilon} \right)^2 nc_j(1 + \delta) =: n^{-1} K(\delta, \varepsilon) U, \quad (3.15)$$

where $K(\delta, \varepsilon) < \infty$ by Assumption A2 (a).

Clearly, from Eq. 3.15 and from Kolmogorov's inequality,

$$\mathbb{P}_i[\phi < \min\{U, \tau_\delta\}] \leq K(\delta, \varepsilon) U / n.$$

Hence, again from Eq. 3.15,

$$\mathbb{E}_i \{ [1 - \rho_{s^*}(U, X_n)]_+ \} \leq n^{-1/2} \sqrt{K(\delta, \varepsilon) U} + n^{-1} K(\delta, \varepsilon) U + \mathbb{P}_i[\tau_\delta < U].$$

Substituting this into Eq. 3.13, it follows that

$$\begin{aligned} & \sum_{l \geq 1} \mathbb{P}(N_n(U) = l - 1) \sum_{m \in \mathbb{Z}} |f_{l,i}^U(m) - f_{l-1,i}^U(m)| \\ & \leq 2 \left\{ n^{-1/2} \sqrt{K(\delta, \varepsilon)U} + n^{-1} K(\delta, \varepsilon)U \right. \\ & \quad \left. + \mathbb{P} \left[\sup_{0 \leq u \leq U} |Z_n(u) - nc| > n\delta \mid Z_n(0) = i \right] \right\}. \end{aligned}$$

But now, for all i such that $|i - nc| \leq n\delta' = n\delta e^{-K_1 U}/2$, the latter probability is of order $O(n^{-1})$, by Lemma 3.1, and hence the final term in Eq. 3.8 is also of order $O(n^{-1/2})$, as required. \square

As a consequence of this theorem, we have the following corollary.

Corollary 3.3 *Under Assumptions A1–A4, for any bounded function f ,*

$$\mathbb{E}_{\Pi_n} \{\nabla f(Z_n)\} = O\left(\frac{1}{\sqrt{n}} \|f\|\right).$$

Proof Immediate, because

$$|\mathbb{E}_{\Pi_n} \{\nabla f(Z_n)\}| \leq 2 \|f\| d_{TV}(\Pi_n, \Pi_n * \delta_1). \quad \square$$

4 Translated Poisson Approximation to the Equilibrium Distribution

We are now able to prove our main theorem. The centred equilibrium distribution of Z_n is $\widehat{\Pi}_n := \Pi_n * \delta_{-\lfloor nc \rfloor}$, and we approximate it by a centred Poisson distribution with similar variance.

Theorem 4.1 *Under Assumptions A1–A5,*

$$d_{TV}(\widehat{\text{Po}}(nv_c), \widehat{\Pi}_n) = O(n^{-\alpha/2}),$$

where $v_c := \sigma^2(c)/\{-2F'(c)\}$.

Proof We follow the recipe outlined in Section 1.1. From Eq. 1.9, we principally need to show that

$$\sup_{g \in \mathcal{G}_v} |\mathbb{E}\{v \nabla g(W + 1) - Wg(W) + \langle v \rangle g(W)\}| = O(n^{-\alpha/2}),$$

for $W := Z_n - \lfloor nc \rfloor$, $v := nv_c$ and $\mathbb{E} := \mathbb{E}_{\Pi_n}$. So, for any $g \in \mathcal{G}_{nv_c}$, write $\tilde{g}(i) := g(i - \lfloor nc \rfloor)$, and set

$$h := h_{n,g}(i) := \begin{cases} 0, & \text{if } i \leq \lfloor nc \rfloor - \lfloor nv_c \rfloor; \\ \sum_{l=\lfloor nc \rfloor - \lfloor nv_c \rfloor}^{i-1} \tilde{g}(l) & \text{if } i > \lfloor nc \rfloor - \lfloor nv_c \rfloor. \end{cases}$$

Note that, for $j \geq 1$, by Assumption A2 (a),

$$\begin{aligned} n\lambda_j(i/n)|h(i+j) - h(i)| &\leq njc_j\|\tilde{g}\| + c_j|i - \lfloor nc \rfloor| \sum_{k=1}^j |g(i+j-k - \lfloor nc \rfloor)| \\ &\leq njc_j\|g\| + jc_j \sup_l |lg(l)| + c_j \sum_{k=1}^j |j-k|\|g\|, \end{aligned}$$

and that a similar bound, with $|j|$ replacing j , is valid for $j \leq -1$. From the definition of \mathcal{G}_{nv_c} in Eqs. 1.6 and 1.7 and from Assumption A2 (a), it thus follows that $(\mathcal{A}_n|h_{n,g})$ is a bounded function, and hence that the function $h_{n,g}$ satisfies condition 1.12; furthermore, since $|h_{n,g}(i)| \leq |i - \lfloor nc \rfloor| + \lfloor nv_c \rfloor$, in view of Eq. 1.7, $h_{n,g}$ is integrable with respect to Π_n , because of Theorem 2.1. Hence it satisfies the conditions of Theorem 1.2, from which we deduce, as in Eq. 1.1, that

$$\mathbb{E}_{\Pi_n}(\mathcal{A}_n h_{n,g})(Z_n) = 0.$$

Applying Lemma 1.1, since $h_{n,g}$ has bounded differences in view of Eq. 1.7, it follows that

$$\begin{aligned} 0 &= \mathbb{E}_{\Pi_n} \left\{ \frac{n}{2} \sigma^2 \left(\frac{Z_n}{n} \right) \nabla \tilde{g}(Z_n) + nF \left(\frac{Z_n}{n} \right) \tilde{g}(Z_n) + E_n(\tilde{g}, Z_n) \right\} \\ &= -F'(c) \mathbb{E}_{\Pi_n} \{ nv_c \nabla \tilde{g}(Z_n) - (Z_n - \lfloor nc \rfloor) \tilde{g}(Z_n) + \langle nv_c \rangle \tilde{g}(Z_n) \} \\ &\quad + \mathbb{E}_{\Pi_n} \{ E'_n(\tilde{g}, Z_n) + E_n(\tilde{g}, Z_n) \}, \end{aligned} \quad (4.1)$$

where E_n is as defined in Eq. 1.3, and

$$\begin{aligned} E'_n(g, i) &:= \frac{n}{2} (\sigma^2(i/n) - \sigma^2(c)) \nabla g(i) + \{ n(F(i/n) - F(c)) - F'(c)(i - \lfloor nc \rfloor) \} g(i) \\ &\quad + F'(c) \langle nv_c \rangle g(i). \end{aligned}$$

The terms involving $E'_n(\tilde{g}, i)$ can be bounded, using Eq. 1.7, as follows. First, using Assumptions A2 (a) and A4,

$$\begin{aligned} \frac{n}{2} |\sigma^2(i/n) - \sigma^2(c)| |\nabla \tilde{g}(i)| &\leq \frac{1}{2nv_c} \|(\sigma^2)'\|_\delta |i - nc| I[|i - nc| \leq n\delta] \\ &\quad + \frac{1}{2v_c} \left(\sum_{j \in \mathbb{Z} \setminus \{0\}} j^2 c_j (1 + |i/n - c|) + \sigma^2(c) \right) \\ &\quad \times I[|i - nc| > n\delta]; \end{aligned} \quad (4.2)$$

and then, under Assumptions A2 (a) and A5,

$$\begin{aligned}
 & |n(F(i/n) - F(c)) - F'(c)(i - \lfloor nc \rfloor) + F'(c)(nv_c)| |\tilde{g}(i)| \\
 &= n|F(i/n) - F(c) - (i/n - c)F'(c)| |\tilde{g}(i)| \\
 &\leq \left(\frac{n}{2} (i/n - c)^2 I[|i/n - c| \leq \delta] \sup_{|z-c| \leq \delta} |F''(z)| \right. \\
 &\quad \left. + n \left\{ (1 + |i/n - c|) \sum_{j \in \mathbb{Z} \setminus \{0\}} |j|c_j + F'(c)|i/n - c| \right\} I[|i - nc| > \delta] \right) \frac{1}{\sqrt{nv_c}}.
 \end{aligned} \tag{4.3}$$

The contribution to Eq. 4.1 from $\mathbb{E}_{\Pi_n}\{E'_n(\tilde{g}, Z_n)\}$ is thus of order

$$\mathbb{E}_{\Pi_n} \{ |z_n - c| + (1 + |z_n - c|) I[|z_n - c| > \delta] + |z_n - c|^2 I[|z_n - c| \leq \delta] \} = O(n^{-1/2}), \tag{4.4}$$

by Theorem 2.1 and Corollaries 2.4 and 2.5. The first term in $E_n(\tilde{g}, i)$ is also bounded in similar fashion: from Assumptions A1, A2 (a) and A4,

$$\frac{n}{2} |F(i/n)| |\nabla \tilde{g}(i)| \leq \frac{1}{2nv_c} \left\{ \|F'\|_\delta |i - nc| + \sum_{j \in \mathbb{Z} \setminus \{0\}} c_j |j| (1 + |i - nc|) I[|i - nc| > \delta] \right\}, \tag{4.5}$$

giving a contribution to $\mathbb{E}_{\Pi_n}\{E_n(\tilde{g}, Z_n)\}$ of the same order. The remaining terms, involving $\nabla^2 \tilde{g}$, need to be treated more carefully.

We examine the first of them in detail, with the treatment of the second being entirely similar. First, if either $|i/n - c| > \delta$ or $j > \sqrt{n}$, it is enough to use the expression in Eq. 1.4 to give

$$|a_j(\tilde{g}, i)| \leq j(j-1) \|\nabla \tilde{g}\| \leq j(j-1)/(nv_c). \tag{4.6}$$

For $|i/n - c| > \delta$, by Assumption A2 (a), this yields the estimate

$$\left| \sum_{j \geq 2} a_j(\tilde{g}, i) n \lambda_j(i/n) \right| I[|i - nc| > \delta] \leq \sum_{j \geq 2} \frac{j(j-1)c_j}{v_c} (1 + |i/n - c|) I[|i - nc| > \delta], \tag{4.7}$$

with the corresponding contribution to $\mathbb{E}_{\Pi_n}\{E_n(\tilde{g}, Z_n)\}$ being of order $O(n^{-1})$, by Theorem 2.1 and Corollary 2.5. Then, for $j > \sqrt{n}$ and $|i/n - c| \leq \delta$, Eq. 4.6 yields

$$\left| \sum_{j > \sqrt{n}} a_j(\tilde{g}, i) n \lambda_j(i/n) \right| \leq \sum_{j > \sqrt{n}} \frac{j(j-1)c_j}{v_c} (1 + \delta) \leq \sum_{j \geq 1} j^{2+\alpha} c_j n^{-\alpha/2} (1 + \delta) / v_c, \tag{4.8}$$

making a contribution of order $O(n^{-\alpha/2})$ to $\mathbb{E}_{\Pi_n}\{E_n(\tilde{g}, Z_n)\}$, again using Assumption A2 (a). In the remaining case, in which $j \leq \sqrt{n}$ and $|i/n - c| \leq \delta$, we use Eq. 1.5, observing first that

$$\begin{aligned} n \nabla^2 \tilde{g}(i + j - k + 1) \lambda_j(i/n) &= n \nabla^2 \tilde{g}(i + j - k + 1) \lambda_j(c) \\ &\quad + n \nabla^2 \tilde{g}(i + j - k + 1) (\lambda_j(i/n) - \lambda_j(c)), \end{aligned} \quad (4.9)$$

the latter expression being bounded by

$$|n \nabla^2 \tilde{g}(i + j - k + 1) (\lambda_j(i/n) - \lambda_j(c))| \leq \frac{2}{v_c} \|\lambda'_j\|_\delta |i/n - c|. \quad (4.10)$$

The corresponding contribution to $\mathbb{E}_{\Pi_n}\{E_n(\tilde{g}, Z_n)\}$ is thus at most

$$\begin{aligned} &\sum_{j=2}^{\lfloor \sqrt{n} \rfloor} (j^3/6) \{ \lambda_j(c) n \sup_l |\mathbb{E}_{\Pi_n} \nabla^2 \tilde{g}(Z_n + l)| + 2v_c^{-1} \|\lambda'_j\|_\delta \mathbb{E}_{\Pi_n} |z_n - c| \} \\ &\leq n^{(1-\alpha)/2} \sum_{j \geq 2} j^{2+\alpha} c_j \left\{ n \sup_l |\mathbb{E}_{\Pi_n} \nabla^2 \tilde{g}(Z_n + l)| + L_1 2v_c^{-1} \mathbb{E}_{\Pi_n} |z_n - c| \right\} \\ &= n^{(1-\alpha)/2} O(n \cdot n^{-3/2} + n^{-1/2}) = O(n^{-\alpha/2}), \end{aligned} \quad (4.11)$$

where we have used Assumptions A2 (a) and A4, and then Corollaries 2.4 and 3.3, and finally Eq. 1.7.

Combining the bounds, and substituting them into Eq. 4.1, it follows that

$$\begin{aligned} &|\mathbb{E}_{\Pi_n} \{nv_c \nabla g(Z_n - \lfloor nc \rfloor) - (Z_n - \lfloor nc \rfloor)g(Z_n - \lfloor nc \rfloor) \\ &\quad + \langle nv_c \rangle g(Z_n - \lfloor nc \rfloor)\}| = O(n^{-\alpha/2}), \end{aligned}$$

uniformly in $g \in \mathcal{G}_{nv_c}$. Again from Corollary 3.3, we also have

$$|nv_c \mathbb{E}_{\Pi_n} \{\nabla g(Z_n - \lfloor nc \rfloor) - \nabla g(Z_n - \lfloor nc \rfloor + 1)\}| = O(n^{-1/2}),$$

for any $g \in \mathcal{G}_{nv_c}$. It thus follows from Eq. 1.9 that

$$d_{TV}(\widehat{\text{Po}}(nv_c), \widehat{\Pi}_n) = O(n^{-\alpha/2} + \mathbb{P}_{\Pi_n}[Z_n - nc < -\lfloor nv_c \rfloor]),$$

and the latter probability is of order $O(n^{-1})$ by Corollary 2.5. This completes the proof. \square

Example Consider an immigration birth and death process Z , with births occurring in groups of more than one individual at a time. The process has transition rates as in Section 1.1, with

$$\lambda_{-1}(z) := dz, \quad \lambda_1(z) := a + bq_1 z \quad \text{and} \quad \lambda_j(z) := bq_j z, \quad j \geq 2,$$

while $\lambda_j(z) := 0$, $j < -1$. Here, b denotes the rate at which birth events occur, and $a > 0$ represents the immigration rate. The quantity q_j denotes the probability that j offspring are born at a birth event, so that $\sum_{j \geq 1} q_j = 1$; we write $m_r := \sum_{j \geq 1} j^r q_j$ for the r 'th moment of this distribution. Then

$$F(z) = a + z(bm_1 - d), \quad \text{and} \quad \sigma^2(z) = a + z(bm_2 + d).$$

Assumption A1 is satisfied if $d > bm_1$, with $c = a/(d - bm_1)$ and $F'(c) = -(d - bm_1)$. Assumption A2(a) is satisfied with $c_j = bq_j \max\{1, c\}$, $j \geq 2$, $c_1 = \max\{bq_1, a + bq_1c\}$, and $c_{-1} = d \max\{1, c\}$, provided that $m_{2+\alpha} < \infty$ for some $0 < \alpha \leq 1$; for Assumption A2(b), simply take $\lambda^0 = a/2$. The other assumptions are immediate.

The quantity v_c appearing in Theorem 4.1 then comes out to be

$$v_c := \frac{a(2d + b(m_2 - m_1))}{2(d - bm_1)^2},$$

and the approximation to the equilibrium distribution of $Z_n - \lfloor nc \rfloor$ is the centred Poisson distribution $\widehat{\text{Po}}(nv_c)$, accurate in total variation to order $O(n^{-\alpha/2})$. Note that, if $b = 0$, then the process becomes a simple immigration death process, whose equilibrium distribution is precisely the Poisson distribution $\text{Po}(na/d) = \text{Po}(nc)$. In this special case, the approximation is in fact exact.

References

- Barbour AD (2003) Equilibrium distributions for Markov population processes. *Adv Appl Probab* 12:591–614
- Barbour AD, Holst L, Janson S (1992) Poisson approximation. Oxford University Press, Oxford
- Ethier SN, Kurtz TG (1986) Markov processes: characterization and convergence. Wiley, New York
- Hamza K, Klebaner FC (1995) Conditions for integrability of Markov chains. *J Appl Probab* 32: 541–547
- Kurtz TG (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes. *J Appl Probab* 7:49–58
- Kurtz TG (1971) Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J Appl Probab* 8:344–356
- Kurtz TG (1981) Approximation of population processes. CBMS-NSF regional conference series in applied mathematics vol 36. SIAM, Philadelphia
- Röllin A (2005) Approximation of sums of conditionally independent random variables by the translated Poisson distribution. *Bernoulli* 11:1115–1128